Technology Arts Sciences Cologne
Faculty of Economics, Business and Law
Prof. Dr. Arrenberg
Room 221, Tel. 39 14
jutta.arrenberg@th-koeln.de

Exercises Quantitative Methods

Worksheet: Goodness-of-fit Test

Exercise 3.1 (*DAX_30.sav*)
DAX 30 (Deutsche Aktien Xchange 30, former Deutscher Aktien-Index 30) is a stock market index consisting of the 30 major German companies trading on the Frankfurt Stock Exchange. It is computed daily between 09:00 and 17:30 Hours CET.

For the first time the index was listed on 30th December in 1987 with the arbitrary index 1 000. For the last years we get the following indices on December 31^{st}:

Year	Dax
1987	1 000
1988	1 328
1989	1 790
1990	1 398
1991	1 578
1992	1 545
1993	2 267
1994	2 107
1995	2 254
1996	2 889
1997	4 250
1998	5 002
1999	6 958
2000	6 434
2001	5 160
2002	2 893
2003	3 965
2004	4 256
2005	5 408
2006	6 597
2007	8 067
2008	4 810
2009	5 957
2010	6 914
2011	5 898
2012	7 612
2013	9 552
2014	9 806

Year	Dax
2015	10 743
2016	11 481
2017	12 918
2018	10 559
2019	13 249
2020	13 719

Source: *Die Süddeutsche Zeitung*

In business financing we assume Normal distribution of the logarithm of the factors; i.e. $\ln \dfrac{\text{DAX}_{\text{Year}}}{\text{DAX}_{\text{Prior Year}}} \sim \mathsf{NV}$. Check this assumption for the last ten factors $\frac{\text{Dax}_{2011}}{\text{Dax}_{2010}}, \ldots, \frac{\text{Dax}_{2020}}{\text{Dax}_{2019}}$ with the Shapiro-Wilk Test.

Descriptives

			Statistic	Std. Error
LN_Faktor	Mean		,0685	,0
	4895% Confidence Interval	Lower Bound	-,0422	
	for Mean	Upper Bound	,1793	
	5% Trimmed Mean		,0732	
	Median		,0788	
	Variance		,024	
	Std. Deviation		,15479	
	Minimum		-,20	
	Maximum		,26	
	Range		,46	
	Interquartile Range		,25	
	Skewness		-0,637	0,687
	Kurtosis		-0,352	1,334

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Signifikanz	Statistic	df	Significance
ln_Faktor	0,192	10	0,200*	0,911	10	0,288

* This is a lower bound of the true significance.

a. Lilliefors Significance Correction

How to compute the factor with SPSS?

1. Put the values of the variable DAX30 in the second row of the next column.
   Denote the new column with “Variable1“
2. Click Transform $\rightarrow$ Compute Variable
3. Target Variable = *factor*
4. Numeric Expression: DAX30/Variable1
5. ok

How to compute Ln(factor) with SPSS?

1. Transform $\rightarrow$ Compute Variable
2. Target Variable = *name*
3. Numeric Expression:
   Select “ln“ of the Function group “Arithmetic“
   Numeric Expression = *ln(factor)*
4. ok

Purpose: Predict probabilities due to the Normal distribution

Problem: The values of a random sample from the standard Normal distribution are lying in the interval $(-\infty; +\infty)$. But the rate of change is an element of the interval $[-1; +\infty)$, because the greatest decrease is -100%. And the factor of change (factor=rate + 1) is an element of the interval $[0; +\infty)$, so Normal distribution is inappropriate for the rate as well for the factor.

Solving of the problem: We consider the random variable $X = \ln(factor)$, so the values of X are lying in the interval $(-\infty; +\infty)$. With the goodness-of-fit test we confirm Normal distribution of X based on the sample of the last seven factors $\frac{\text{index}_{2012}}{\text{index}_{2011}}, \frac{\text{index}_{2013}}{\text{index}_{2012}}, \ldots, \frac{\text{index}_{2018}}{\text{index}_{2017}}$, i.e. $X \approx \mathsf{N}(\mu = 0.0832, \sigma = 0.15065)$

We invest 10 000 € in a German stock for one year.

a) What is the probability of the event, that the loss in the next year does not exceed 1 000 €?

b) Value at risk? What minimum charge would we expect with probability 95% in the next year?

Solution

a) 2021: investment of 10 000 €
2022: loss less than 1 000 €
factor=Value 2022/Value 2021=$\frac{9\,000}{10\,000} = 0.9$
$\ln(factor) = \ln(0.9) = -0.1054$

$P(X > -0.1054) = 1 - P(X \leq -0.1054) = 1 - \phi\left(\frac{-0.1054 - 0.0685}{0.15479}\right) =$
$1 - \phi(-1.123458) = 1 - 0.131 = 0.869 = 86.9\%$
With the probability of 86.9% the loss in the next year will not exceed 1 000 €.

b) $0.05 = P(X \leq x) = \phi\left(\frac{x - 0.0685}{0.15479}\right)$
$-1.6449 = \frac{x - 0.0685}{0.15479}$
$x = 0.0685 - 1.6449 \cdot 0.15479 = -0.3231141 = \ln(factor)$
$e^{-0.3231141} = factor = 0.724$
$10\,000 \cdot 0.724 = 7\,240$ €
$10\,000 - 7\,240 = 2\,760$ €
that is with probability of 95% the loss will not exceed 2 760 €.

Technology Arts Sciences Cologne
Faculty of Economics, Business and Law
Prof. Dr. Arrenberg
Room 221, Tel. 39 14
jutta.arrenberg@th-koeln.de

Exercises Quantitative Methods

Worksheet: Goodness-of-fit Test

Exercise 3.2 (Berenson et al. page 268)
The data in the file *chicken.sav* contains the total fat, in grams per serving, for a sample of 20 chicken sandwiches from fast-food chains. The data are as follows:

7	8	4	5	16	20	20	24	19	30
23	30	25	19	29	29	30	30	40	56

Please decide whether the data appear to be approximately normally distributed by running a test.

Exercise 3.3 (Berenson et al. page 268)
The following data, stored in the file *electricity.sav*, represents the electricity costs in dollars during July 2007 for a random sample of 50 two-bedroom apartments in a large city:

96	171	202	178	147	102	153	197	127	82
157	185	90	116	172	111	148	213	130	165
141	149	206	175	123	128	144	168	109	167
95	163	150	154	130	143	187	166	139	149
108	119	183	151	114	135	191	137	129	158

Decide whether the data appear to be approximately normally distributed by running a test.

Solution exercise 3.2
X=total fat of a chicken sandwich
$n = 20$ chicken sandwiches
p-value Lilliefors test $= 0.053$
i.e. normal distribution

p-value Shapiro-Wilk test$= 0.150$
i.e. normal distribution

Solution exercise 3.3
X=electricity costs in dollars for a two-bedroom apartment
$n = 50$ apartments
p-value Lilliefors test ≥ 0.2
i.e. normal distribution

p-value Shapiro-Wilk test$= 0.941$
i.e. normal distribution